Smart Blockchain Badges for Data Science Education

Conference or Workshop Item

Smart Blockchain Badges for Data Science Education

Alexander Mikroyannidis, John Domingue, Michelle Bachler and Kevin Quick
Knowledge Media Institute, The Open University
Milton Keynes, United Kingdom
{Alexander.Mikroyannidis, John.Domingue, Michelle.Bachler, Kevin.Quick}@open.ac.uk

***Abstract*—Blockchain technology has the potential to revolutionise education in a number of ways. In this paper, we explore the applications of Smart Blockchain Badges on data science education. In particular, we investigate how Smart Blockchain Badges can support learners that want to advance their careers in data science, by offering them personalised recommendations based on their learning achievements. This work aims at enhancing data science accreditation by introducing a robust system based on the Blockchain technology. Learners will benefit from a sophisticated, open and transparent accreditation system, as well as from receiving job recommendations that match their skills and can potentially progress their careers. As a result, this work contributes towards closing the data science skills gap by linking data science education to the industry.**

Keywords—Blockchain; Data Science; Open Badges; Smart Blockchain Badges; Accreditation.

I. INTRODUCTION

The 'Age of Data' is currently thriving, with data being produced from all industries at a phenomenal rate that introduces numerous challenges regarding the collection, storage and analysis of this data. Declared by Harvard Business Review as the "sexiest job of the 21st century" [1], data science skills have become a key asset in any organisation confronted with the daunting challenge of making sense of information that comes in varieties and volumes never encountered before.

However, as the amount of available data continues to increase, so does the demand for professionals that have the necessary skills to manage and manipulate this data. A McKinsey study estimated already in 2011 that the United States will soon require 60 percent more graduates able to handle large amounts of data as part of their daily jobs [2]. With an economy of comparable size (by GDP) and growth prospects, Europe will most likely be confronted with a similar talent shortage of hundreds of thousands of qualified data scientists, and an even greater need of executives and support staff with basic data literacy. The number of job descriptions and an increasing demand in higher-education programs and professional training confirm this trend [3], with some EU countries forecasting an increase of almost 100 percent in the demand for data science positions in less than a decade [4].

The European Data Science Academy (EDSA)[1] is an initiative for bridging the data science skills gap across Europe and beyond, as well as for training a new generation of world-leading data scientists. The EDSA project offers interactive tools for exploring the current demand for data science jobs and skills. Additionally, EDSA offers tools for finding learning resources and building personalised learning pathways towards acquiring the skills that are currently in demand.

In order to facilitate accreditation and offer personalised recommendations to learners that study data science courses, we are currently in the process of developing and deploying an infrastructure for Smart Blockchain Badges. Learners that study various data science subjects are earning badges upon reaching certain milestones in their studies, e.g. completing part of a course or a whole course. These badges are stored on the Blockchain and include data about the key skills that learners have acquired upon obtaining these badges. As learners continue to earn badges, they start receiving automated recommendations for the latest job offers that match their skills. Datasets of current job offers and their associated skills are being harvested by the EDSA tools from various sources. These datasets are then placed in contracts on the Blockchain, which are used for matching jobs with a learner's badge-based skills. In this way, the awarded badges are smart, in the sense that they are being used to offer recommendations to learners. We have based our approach on the Blockchain technology, as the latter offers the means for resilient management of distributed data, thus enabling us to enhance the flexibility and scalability in the ways accreditation is defined and granted. Additionally, the Blockchain's immutable ledger infrastructure ensures the validity of the awarded accreditation and eliminates the risk of fabricated qualifications.

The remainder of this paper is organised as follows. Section II introduces the interactive tools developed by the EDSA project for finding data science job offerings and training. Section III introduces Open Badges, followed by Section IV, which presents our approach for generating and using Smart Blockchain Badges in data science education, as well as potential applications beyond data science education. Finally, the paper is concluded and the next steps of this work are outlined.

[1] http://edsa-project.eu

II. THE EDSA DASHBOARD

Linking the demand for data science skills with the supply of learning resources that offer these skills is crucial for bridging the data science skills gap. Towards this goal, the EDSA project has developed an interactive dashboard[2] (shown in Figure 1) that enables its users to explore both the current data science skills demand and supply [5]. Additionally, users are supported in building personalised learning pathways, consisting of courses and learning materials that will help them reach their learning goals.

Automated tools have been deployed for extracting data about job posts and news articles in order to present the current state of the European data science landscape. Job data from search services that aggregate job posts are harvested by the EDSA dashboard and automatically annotated with the relevant data science skills. Some of the job aggregators harvested include Adzuna,[3] Trovit,[4] Jooble,[5] Indeed[6] and XING[7]. A total of approximately 3.3 million job posts have been harvested and made available via the dashboard, covering a period of 18 months. This dataset is being constantly updated as new job posts are advertised in the aforementioned job aggregators.

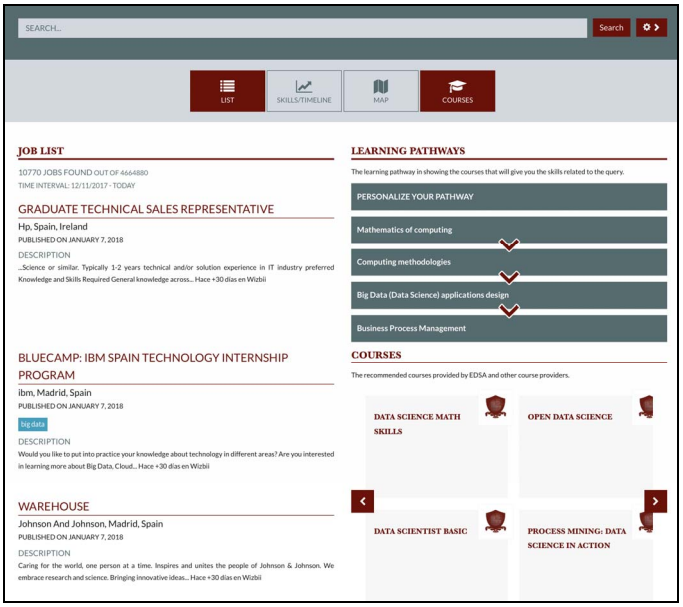

Fig. 1. The EDSA dashboard – exploring the current demand for data science skills and the supply of training.

In summary, the EDSA dashboard enables users to:

- View the current demand for data science jobs and skills across Europe.
- Filter demand by required skills and region.
- View trends and statistics regarding data science jobs and skills for a given timeframe.
- Explore the current supply of courses and learning materials that will help them acquire certain skills.
- Build personalised learning pathways towards acquiring certain skills.

III. OPEN BADGES

Open Badges[8] allow for detailed recording of accreditation in digital form from both formal and informal learning contexts. Open Badges were initiated by the Mozilla Foundation in 2010 [6] as a way of providing a verifiable digital recognition of learning across a wide variety of contexts, including:

- ***Capturing a learning path*** – in essence breaking up a single large qualification, such as a degree, into constituent parts giving a detailed account as to what has been achieved. The learning path may cross institutions.
- ***Achievement signalling*** – indicating to the outside world specific skills or achievements. For example, enabling recruiters to identify suitable candidates.
- ***Motivation*** – through intrinsic feedback encouraging continued engagement and retention. Additionally, badges can enable awareness of or grant specific privileges.
- ***Innovation and flexibility*** – enabling the capture of skills which may be missed or ignored within formal accreditation and newer emerging skills for example, related to particular forms of digital literacy. Badges provide a flexible channel to recognise new or currently unrecognised skills.
- ***Identity/reputation building*** – badges can promote identity and reputation within learning and peer communities. Any existing individual and aggregate identity and reputation can be made explicit and portable across communities and peer groups.
- ***Community building/kinship*** – membership of a community can be signalled enabling peers with similar interest to be found or potential mentors or teachers. Badges are a mechanism for providing social capital and the formalisation of camaraderie and communities of practice.

Since 2010, millions of Open Badges have been awarded[8] and have been taken up by a number of organisations including the Clinton Global Initiative[9] and NASA.[10] Although in extensive use, a number of problems have been articulated with the use of Open Badges. Belshaw [7] notes that often complaints are made on how the value of a badge

[2] http://edsa-project.eu/resources/dashboard/
[3] https://www.adzuna.com/
[4] https://www.trovit.com/
[5] https://jooble.org
[6] https://indeed.com/
[7] https://www.xing.com/jobs

[8] https://openbadges.org/about/
[9] https://www.macfound.org/press/press-releases/better-futures-2-million-americans-through-open-badges/
[10] https://www.nasa.gov/offices/education/programs/national/dln/special/DigitalBadges.html

can be judged. With an Open Badges infrastructure there are no gatekeepers, meaning that anyone is allowed to issue a badge for anything. In their assessment of 29 badge development efforts, Hickey et al. [8] found that unsuccessful projects were hindered by problems of interoperability and integration between badging systems and institutional platforms. On the other hand, successful badge deployments layered badges on top of existing content and infrastructure, were tied to public student ePortfolios and used a mix of automated and human expert issued awards. They also found key to success were the embedding of badges in more social learning forms and ensuring that the badges contain unique non-redundant information. The most promising route for badges, they argue, is to link to formal externally recognised certificates whilst adding additional claims and evidence.

IV. Smart Blockchain Badges

The Blockchain offers a publicly shared immutable ledger, a technology that enables the secure and resilient management of distributed data in combination with data analytics techniques that add scale and flexibility to the way levels of qualifications are defined and granted.

Smart Blockchain Badges are dynamic records of accreditation that follow the same principles as Open Badges and offer the same benefits in recording accreditation as outlined in the previous section of this paper. However, the key difference and novelty of Smart Blockchain Badges lies in their dynamic features. For example, apart from just recording a learning achievement, a Smart Blockchain Badge can also offer job or course recommendations as described in the next section of this paper. These dynamic features are implemented as Smart Blockchain Contracts, which can be defined as "automatable and enforceable agreements" [9]. Smart Blockchain Contracts constitute one of the main features of current Blockchain platforms, such as Ethereum [10].

The following sections describe our approach for generating and using Smart Blockchain Badges. First, we describe the data science education scenario that we have implemented, followed by the implementation details. Finally, we outline the applications of this approach within and beyond data science education.

A. *Data science education scenario*

Our learner, Michelle, is looking to expand her knowledge on data science, and has thus enrolled to a number of courses offered online, including Massive Open Online Courses (MOOCs), Open Educational Resources (OERs), as well as online paid courses. Each time she completes a course, she earns a Smart Blockchain Badge. These badges are stored on the Blockchain and include data about the skills that Michelle has acquired.

After studying for several months, Michelle has mastered some basic data science skills, including various computer science topics such as databases. Based on these skills, the Smart Blockchain Badges generate recommendations about jobs that may be suitable for Michelle. As shown in Figure 2, Michelle is getting recommendations about jobs that fully match her skills, as well as about jobs that match her skills partially. Michelle may also filter the job recommendations according to the location of the job, via the map view shown in Figure 3.

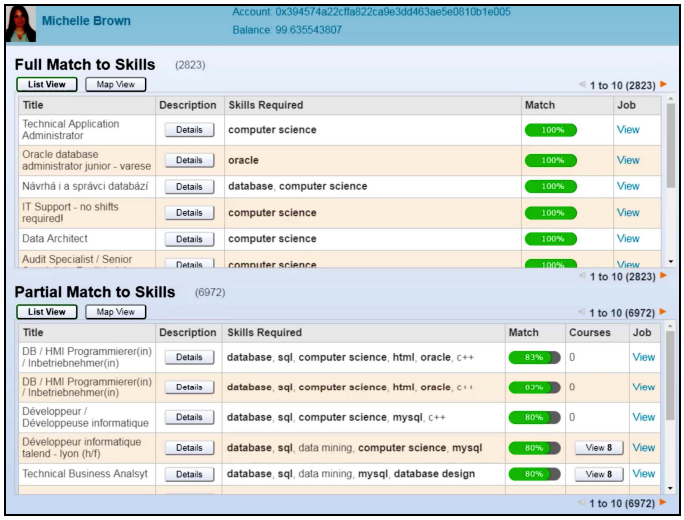

Fig. 2. Job recommendations based on full and partial matches to acquired skills.

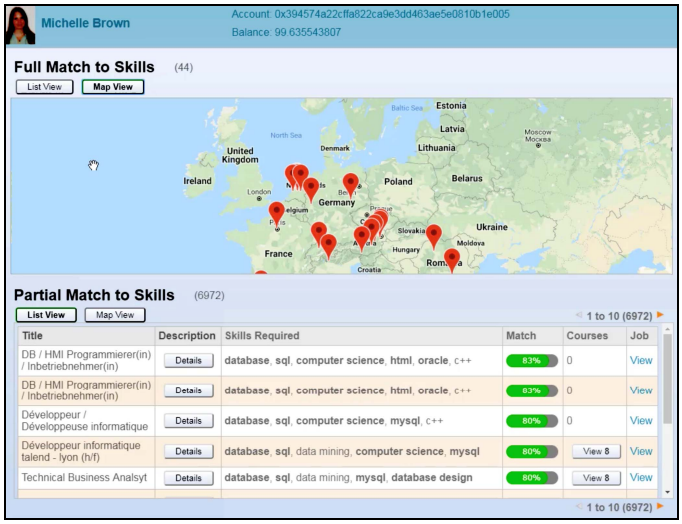

Fig. 3. Map view of job recommendations.

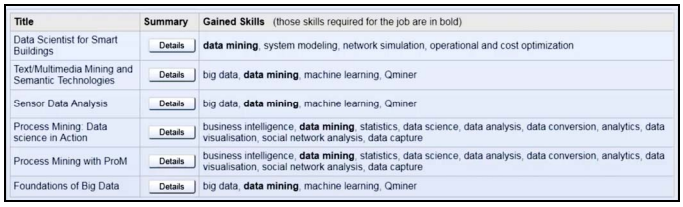

Fig. 4. Course recommendations for obtaining the additional skills assoicated with a particluar job.

Michelle is interested in one of the jobs that matches her skills partially. She then receives recommendations about courses that will give her the additional skills required for this

job (Figure 4). By using this Blockchain-based infrastructure to support her in her studies, Michelle can adopt a more efficient and targeted approach to learning, towards achieving her desired career trajectory.

B. *Implementation*

Figure 5 outlines our implementation process, which is based on the use of Smart Contracts for the Ethereum Blockchain platform. In particular, we have generated the following types of Smart Contracts:

- ***Badge contract*** – it stores the details of a badge that can be awarded upon completion of a course. It contains data like title, description, issuer, criteria for awarding the badge, etc.
- ***SkillsAware contract*** – it is used to store a unique list of all the skills that are required to gain badges and also all the skills as required by jobs that are harvested from the EDSA dashboard.
- ***JobPosting contract*** – it holds the details of a particular job post that has been harvested from the EDSA dashboard, e.g. title, description, country, organisation, location, etc.
- ***JobStore contract*** – it maps JobPosting contracts to skills. It contains a list of pointers to map all the JobPosting contracts (i.e. individual jobs) to the corresponding skills held in the SkillsAware contracts.

In order to generate the job recommendations shown in Figures 2 and 3, we aggregate the skills from Badge contracts and request matching jobs from the JobStore contract. For those jobs that there is a partial skills match, we request suitable courses from the EDSA dashboard in order to offer the recommendations shown in Figure 4. The implementation process and the associated scenario are demonstrated in a short movie available online.[11]

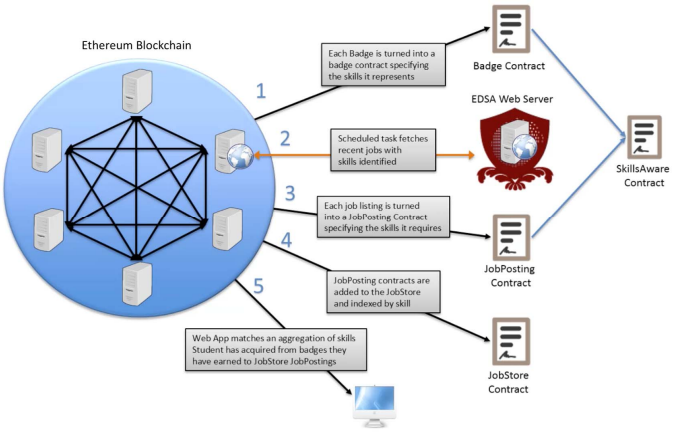

Fig. 5. The process of generating Smart Blockchain Badges by matching the learner's skills with job offerings.

C. *Applications*

Smart Blockchain Badges bear a number of significant benefits to learners, by supporting them in shaping their learning path in order to achieve their career goals. More specifically, we foresee the following applications of Smart Blockchain Badges in different aspects of education:

- ***Job hunting*** – as described in our scenario, Smart Blockchain Badges can be used to link accreditations to jobs through the matching of badges to job profiles.
- ***Promotion*** – similarly to job hunting, Smart Blockchain Badges can be used to match acquired badges to organisational roles for internal promotion.
- ***Mandatory CPD training*** – workers in a variety of professions are obliged to take CPD training annually in order to be able to continue to practice. For example, solicitors in the UK are required to plan suitable learning activities, complete and evaluate their learning, record how the process was carried out and make an annual declaration on their overall actions. Other professions such as banking and healthcare have similar obligations [11]. Smart Blockchain Badges, coupled to a reasoning system based on Linked Data [12], will enable compliance with regulatory CPD training to be checked by relevant authorities and non-compliance to be automatically signalled.
- ***Networking*** – can be achieved using inference-based matches of acquired skills as represented by Smart Blockchain Badges. Networking may be with peers within a specific sector, by general seniority, by geography and also with workers who are less or more senior for mentor-based relationships.
- ***Courses*** – can be automatically recommended based on the gap between current skills and desired jobs (with higher salary or status). Recommendations may be for a single course or a small set of courses. This has also been demonstrated in our data science education scenario.
- ***Reconfiguration*** – with the right granularity, badges can be reconfigured to align with sector specific qualifications. For example, a combination of computing badges aligned to a company specific qualification.

V. Conclusions and Future Work

This work is addressing the data science skills gap by following a Blockchain-driven approach in linking data science education to the industry. Smart Blockchain Badges are employed in order to support learners towards accomplishing their career goals. In our next steps, we will be deploying Smart Blockchain Badges for use by real learners and will be evaluating their usefulness for lifelong learning. We will also be extending this work beyond the data science domain, by deploying it in popular general-purpose online learning platforms, such as the FutureLearn[12] MOOC platform and the OpenLearn[13] OER repository.

[11] http://blockchain7.kmi.open.ac.uk/movies/movies/course-jobs2.mp4

[12] https://www.futurelearn.com

[13] http://www.open.edu/openlearn/

ACKNOWLEDGMENT

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 643937 (EDSA) and No 688095 (SlideWiki).

REFERENCES

[1] T. H. Davenport and D. Patil, "Data scientist: The sexiest job of the 21st century," *Harvard Business Review*, 2012.

[2] M. James, C. Michael, B. Brad, B. Jacques, D. Richard, R. Charles, and H. Angela, "Big data: The next frontier for innovation, competition, and productivity," *The McKinsey Global Institute*, 2011.

[3] B. Glick, "Government calls for more data scientists in the UK," http://www.computerweekly.com/news/2240208220/Government-calls-for-more-data-scientists-in-the-UK, Computer Weekly, 2013.

[4] B. McKenna, "Demand for big data IT workers to double by 2017, says eSkills," http://www.computerweekly.com/news/2240174273/Demand-for-big-data-IT-workers-to-double-by-2017-says-eSkills, Computer Weekly, 2012.

[5] A.-S. Dadzie, E. Sibarani, I. Novalija, and S. Scerri, "Structuring visual exploratory analysis of skill demand," *Journal of Web Semantics*, 2017.

[6] Mozilla Foundation, Peer 2 Peer University, and The MacArthur Foundation, "Open Badges for Lifelong Learning," https://wiki.mozilla.org/images/5/59/OpenBadges-Working-Paper_012312.pdf, 2010.

[7] D. Belshaw, "The three biggest (perceived) problems with Open Badges [Blog post]," https://dougbelshaw.com/blog/2015/04/16/three-biggest-problems-with-badges/, 2015.

[8] D. T. Hickey, J. Willis, and J. Quick, "Where badges work better," *Educause Learning Initiative ELI*, 2015.

[9] C. D. Clack, V. A. Bakshi, and L. Braine, "Smart contract templates: foundations, design landscape and research directions," *arXiv preprint arXiv:1608.00771*, 2016.

[10] V. Buterin, "Ethereum white paper," in *GitHub repository* https://github.com/ethereum/wiki/wiki/White-Paper, 2013.

[11] H. Thomas and T. Qiu, "Continuing professional development: Accountability, autonomy, efficiency and equity in five professions," *British Journal of Educational Studies*, vol. 61, pp. 161-186, 2013.

[12] T. Berners-Lee, "Linked Data - Design Issues," http://www.w3.org/DesignIssues/LinkedData.html, 2006.